



# The Marshal8e6 Library

By Mark Parker | Product Manager, Marshal8e6

## Introduction

The Marshal8e6 library is the collection of categorized URLs, Wildcard URLs, Patterns, URL Keywords and Search Engine Keywords. The library is used in the Marshal8e6 filtering products as a means of categorizing Internet requests. With over 13 years in development, and hundreds of millions of identified web pages, the Marshal8e6 Library is one of the oldest and most comprehensive filtering libraries available today.

## The URL Library

The URL Library consists of millions of URLs divided into 100+ categories. For example [www.google.com](http://www.google.com) is categorized as Search Engines while [mail.google.com](mailto:mail.google.com) is categorized as Web-Based Email. Using Marshal8e6 filtering products customers can choose to log, block, warn, or allow Internet requests based upon the categorization.

## The Wildcard URL Library

When adding URLs to the Marshal8e6 Library, much more protection can be gained by using a wildcard. For example, let's use a fictitious adult site called [yuiririp.com](http://yuiririp.com). Since the domain itself is targeted towards an adult site, the administrator of the site could choose to use sub-domains to separate portions of the site (ex. [women.yuiririp.com](http://women.yuiririp.com), [men.yuiririp.com](http://men.yuiririp.com), [blondes.yuiririp.com](http://blondes.yuiririp.com) etc.) This means that there could be hundreds, or thousands of different sub-domains. Using a wildcard entry of [\\*.yuiririp.com](http://*.yuiririp.com) every possible current sub-domain, and all future sub-domains will be covered. This can also be useful for sites such as [espn.com](http://espn.com), [mlb.espn.com](http://mlb.espn.com), [nfl.espn.com](http://nfl.espn.com), [nba.espn.com](http://nba.espn.com) can all be covered with [\\*.espn.com](http://*.espn.com).

## The Patterns Library

The Patterns Library is a series of patterns designed to provide identification and control of Internet based applications. The Patterns Library provides zero-day protection by identifying the traffic type dynamically regardless of which web page, or website, it is published on. The Patterns Library currently consists of patterns for:

- Web-Based Proxies
- Instant Messaging
- Peer-to-Peer
- Streaming Media
- Remote Desktop
- Online Games

## The URL Keyword Library

The URL Keyword Library consists of character strings which may appear in a URL. For example, if the administrator wanted to block any URLs with the string of character "asdfghjk" this can be done by adding that string of characters to the URL Keyword Library. The URL Keyword Library has a small set of URLs in select categories, such as the strong "proxy" exists in the Web-Based Proxy category. By default the URL Keyword feature is disabled in the Marshal8e6 products, however the administrator can enable this feature.

## The Search Engine Keyword Library

The Search Engine Keyword library is similar to the URL Keyword Library in that it provides the ability to specifically categorize the searches that users perform via search engines. For example, when a user searches for "Myspace Bypass" that can be categorized as "Web-Based Proxies" since the user's intent is to find a Web-Based Proxy.

## User Customization of the Marshal8e6 Library

Users can customize the Marshal8e6 Library on their own Marshal8e6 products. With the exception of the Patterns Library, all portions of the Marshal8e6 Library can be customized through the addition and/or deletion of items in the Library. The customer can also create custom categories that can contain any, or a combination of URLs, Wildcard URLs, URL Keywords, and Search Engine Keywords.

## How the Marshal8e6 Library is maintained by Marshal8e6 TRACELabs

There are six key components that go into keeping the Marshal8e6 Library up to date, day in and day out. These components come together in Marshal8e6 TRACELabs. TRACELabs is a collection of highly skilled Internet and Internet threat experts that work around the clock on multiple continents to ensure that the Marshal8e6 Library is kept up to date with correct categorizations and protections from Internet threats such as Malicious Code, Viruses, Spyware, Blended Threats, Phishing and the latest application signatures.

The five key components that go into the Marshal8e6 Library are Customer Feedback Module, TRACENet, SpamCensor, Marshal8e6 Web Crawlers, 3rd Party Feeds, and human verifiers.

## Customer Feedback Module (CFM)

CFM is feed that comes to TRACELabs from Marshal8e6 Internet Filtering products to TRACELabs. This feed consists of the most frequently visited uncategorized URLs from each customer site (customers who opt to disable the CFM feature on the filtering products will not send data back to TRACELabs, but will still benefit from the collective data which CFM provides. CFM ensures that uncategorized URLs that users visit most often are categorized.

## TRACEnet/SPAMCensor

TRACEnet and SPAMCensor is a feed from Marshal8e6 Mail Filtering products to TRACELabs as well as a collection mail honey pots designed to capture in the wild SPAM, malicious emails and blended email/web threats. Data from these feeds will result in mail and Internet Filtering list changes. For example, TRACEnet will detect an email attack that consists of a SPAM message with links to a site with malicious code (a blended threat.) TRACELabs will identify this threat and update the URL or Wildcard Library (as applicable.) In addition, domains that are consistently associated with blended threats and/or SPAM will be added to the Bad Reputation Domains.

## About Marshal8e6

Marshal8e6 is a global provider of Secure Internet Gateway products for organizations of all sizes. Marshal8e6 is the only security company capable of delivering comprehensive content security across multiple delivery platforms, including software, appliances and Software-as-a-service (SaaS). The company's complete security portfolio delivers the tools necessary to manage and secure email, Web and the endpoint as well as protect against data leakage. Today, more than 16 million end users in more than 20,000 companies in 96 countries rely on Marshal8e6 solutions to protect their businesses at the email and Web gateway.



### Corporate Headquarters Marshal8e6

828 West Taft Avenue  
Orange, CA 92865  
United States

Phone: +1 (714) 282-6111  
Fax: +1 (714) 282-6116

### International Headquarters Marshal8e6

Renaissance 2200  
Basing View, Basingstoke  
Hampshire RG21 4EQ  
United Kingdom

Phone: +44 (0) 1256 848 080  
Fax: +44 (0) 1256 848 060

### Asia-Pacific Marshal8e6

Suite 1, Level 1, Building C  
Millennium Center  
600 Great South Road  
Auckland, New Zealand

Phone: +64 (0) 9 984 5700  
Fax: +64 (0) 9 984 5720